Neural Network Model of a Second Stage L-band Amplifier using Experimental Training Sets

Hamed Rabbani, Kaboko Jean-Jacques Monga, Sophie LaRochelle and Leslie A. Rusch
ECE Department, Center for Optics, Photonics and Lasers, Université Laval, Québec, Canada
*rusch@gel.ulaval.ca

Abstract: Using experimental measurements with high-power input signals, we train a neural network model of the second stage of an L-band amplifier. With the model, we jointly optimize amplifier gain and noise figure (alternately gain flatness). © 2023 The Author(s)

1. Introduction

Erbium doped fiber amplifiers (EDFAs) play a critical role in compensating signal loss and enhancing the efficiency of optical links. A better amplifier design can reduce network power margins and improve system performance. Physics-based EDFA models work well for the C-band, however, they are less accurate in the extended L-band. In this spectral region, extracting precise Giles parameters is difficult due to the low emission cross-section and the presence of excited state absorption [1]. Previously we addressed modeling of a single stage L-band amplifier with an innovative application of neural network (NN). We exploited experimental data sets to predict the behavior of a single-stage EDFA [2]. The very accurate NN-based predictions found with fast computation time are excellent tools for amplifier design and optimization. [2,3].

Compared to a single stage amplifier, a two-stage amplifier can deliver significantly higher gain while preserving a relatively low noise figure (NF). Multi-stage amplifiers are particularly challenging to optimize, whether targeting spectral gain, NF, and/or power efficiency. While the first amplifier stage may see a flat signal profile, the second stage will see both higher powers and less uniform power across wavelength. We extend our previous modeling of a single (or first) stage L-band amplifier. We modify our experimental setup to enrich the diversity of the training set required for a second stage. We evaluate the accuracy of model predictions. Using the NN model, we present an example of an amplifier design with two distinct performance criteria.

2. Experimental Approach and Neural Network Modeling

2.1. Collection of Experimental Training Set

Figure 1 shows the two-stage experimental setup we use to collect training sets. The first stage provides the second stage with sufficient signal power to create a meaningful training set. Both stages use the same erbium doped fiber (EDF) that was designed and produced in our lab to cover the extended L-band [4].

The input to Stage 1 is an L-band laser source with cumulative 0.5 dBm over seven wavelengths ranging from 1575 to 1626 nm (i.e., -7.95 dBm per wavelength). The Stage 1 is held constant throughout the data collection; the output spectra is given as an inset, with signal power per wavelength varying from 5.29 dBm to 15.23 dBm.

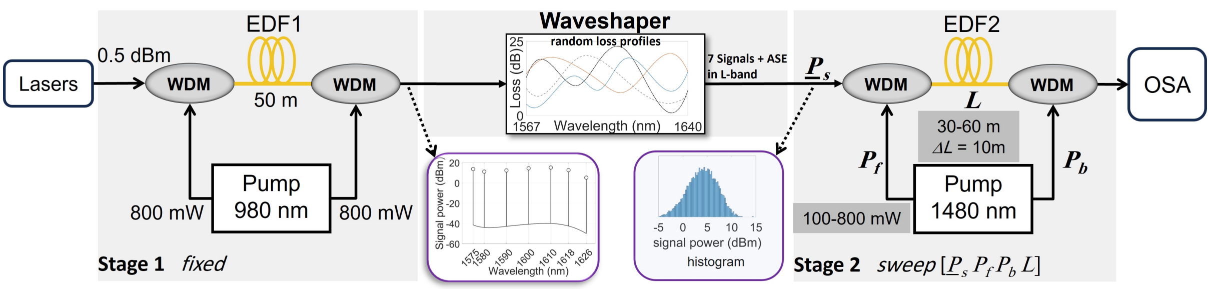


Fig. 1: Experimental setup for capturing the training and test sets (OSA: optical spectrum analyzer). The waveshaper box has four examples from the random loss profiles we applied; insets are for fixed output spectrum of Stage 1, and a histogram of total power input to Stage 2 when using the random loss profiles.

The first stage is followed by a programmable filter covering 1567 to 1640 nm, i.e., a waveshaper (Finisar 1000A/XL, 6 dB insertion loss). For the 0.16 nm band centered on each channel, we create random attenuation levels. We take independent random variables uniformly distributed between 0 and 25 dB for the attenuation at each wavelength. In this way we produce a variety of feasible input signal power profiles, $\underline{P_s}$. Note that the seven random attenuation levels are smoothed using spline interpolation, producing waveshaper profiles such as those shown in the waveshaper box in Fig. 1. The second stage sees an input of seven signals as well as L-band amplified spontaneous emission (ASE).

Stage 2 in Fig. 1 has an EDF with length L. We pump bidirectionally with two 1480 nm laser diodes, at powers P_f and P_b. We used 1480 nm pumps for increased inversion and amplification efficiency while keeping NF low. We sweep the input vector $[\underline{P_s}\ P_f\ P_b\ L]$ to gather the training set. We can sweep $[\underline{P_s}\ P_f\ P_b]$ with automated control of the waveshaper and pumps. However, the preparation of different fiber lengths $[L]$ can only be done by manually cutting the fiber. We collect data at four fiber lengths (30, 40, 50, and 60 m). An optical spectrum analyzer measures output spectra of Stage 2, from which we calculate the gain and NF at each wavelength. The $[P_f\ P_b\ L]$ ranges are noted in Fig. 1. To complement the training set, we gather an additional test set, by randomly selecting vectors $[\underline{P_s}\ P_f\ P_b]$ within training set measurement ranges. We use fiber lengths of 35, 45, and 55 m for the test set.

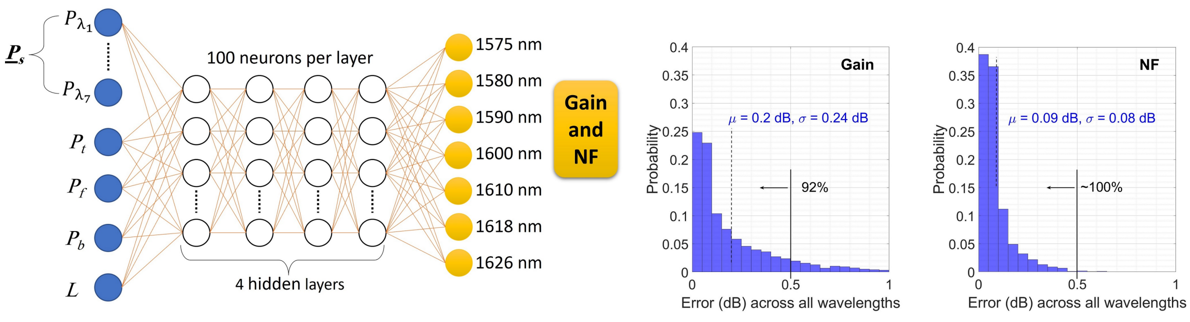
2.2. *NN Structure and Prediction Accuracy*

We collected a training set of 15000 vectors $[\underline{P_s}\ P_f\ P_b\ L]$ over a one week measurement campaign. We train a feed forward neural network (FFNN) with a structure given in Fig. 2a; it is similar to the one in [2]. We have four hidden layers with an equal number of neurons per layer, and scaled exponential linear unit (SELU) activation functions. As shown in Fig. 2a, the current model has a vector of signal power input, whereas in the previous model the signal input was only a scalar of total signal power (i.e., all wavelengths had equal power). Previously we used separate neural networks for predicting gain and NF. Now we use a single FFNN to capitalize on correlations between gain and NF. The error in gain and NF predictions are equally weighted during training.

The test set had N_t =1500 vectors. We define the absolute error in dB between predicted gain (or NF), $\hat{Y}(\lambda)$, and true gain (or NF), $Y(\lambda)$, per

$$\text{Error (dB), all wavelengths} = \left\{ \left| \hat{Y}_j(\lambda_i) - Y_j(\lambda_i) \right|, i = 1,2,\ldots,7 \right\}_{j=1}^{N_t} \tag{1}$$

The histograms in Fig. 2b,c of the absolute error include annotations for the mean, μ, and the standard deviation, σ. For gain in Fig. 2b, 92% of errors are below 0.5 dB. For noise figure in Fig. 2c, the histogram is even more peaked near zero and nearly every error is below 0.5 dB. We expect greater error in gain than noise figure, as gain is more sensitive to variations in fiber length and inversion level. This low prediction error comes along with computations that are one to two orders of magnitude faster than the Runge-Kutta calculations required in physics-based simulation models.



(a) (b) (c)

Fig. 2: a) The FFNN structure, and histograms of the absolute error in dB for b) gain, and c) NF prediction.

3. Jointly Optimizing for Large Gain and Small Ripple (or Noise Figure)

Using our NN model, we can exploit the accuracy and computational efficiency to analyze performance trade-offs across a large parameter space. The model can attack complex designs or interdependent performance metrics. In the following, we examine a simple two-stage amplifier design problem. In this example, we fix the amplifiers and optimize the mid-stage band pass filter, while attacking two different performance targets. That is, we assume that both stages in the amplifier have fixed parameters $[P_f\ P_b\ L]$. In the first stage, the input signal powers are flat and fixed and we use the values specified in Fig. 1. In the second stage $P_f = P_b = 800$ mW and $L = 60$ m, while $\underline{P_s}$ varies according to the setting of the mid-stage band pass filter, e.g., a waveshaper.

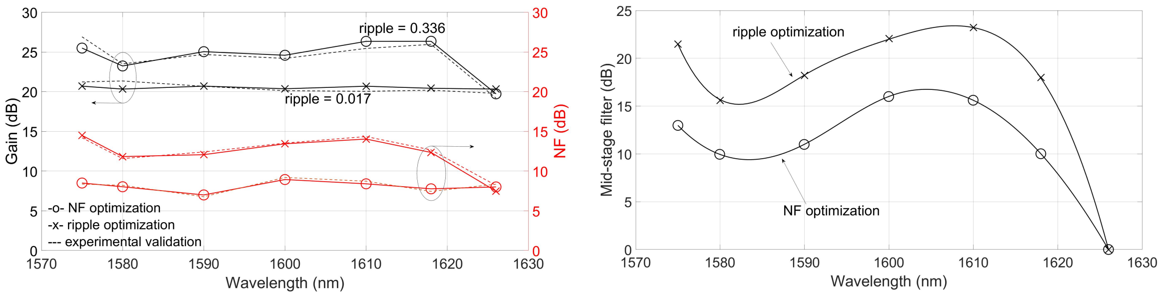


Fig. 3: a) Gain (solid lines) and NF (dashed lines) vs. wavelength for two optimizations, and b) corresponding mid-stage band pass filters for the two optimizations.

We utilize the trained NN model and particle swarm optimization (PSO) technique to perform joint optimization. One optimization is maximizing the gain on the worst-case wavelength, i.e., we maximize the minimum gain. The joint optimization can be while minimizing noise figure or gain ripple. We define $(G_{max} - G_{min})/G_{min}$ over the spectral range from 1575 to 1626 nm to be the gain ripple; the gain is the total gain over the two stages. With PSO we investigate the parameter space through the evolution of a set of solutions. Compared to exhaustive search, we can explore a more extensive parameter space by exploiting PSO computational efficiency. Each PSO particle is a mid-stage band pass filter profile of $[W_{\lambda_1}\ W_{\lambda_2}\ ...W_{\lambda_7}]$. Each PSO particle is initialized randomly within [0, 25] dB, and then particles adjust their velocity based on their individual best objective function value as well as the swarm global best result. We set cognitive coefficient $c_1 = 0.5$, social coefficient $c_2 = 0.5$, and inertia coefficient $w = 0.5$. The optimization process terminates once the difference between the last two iterations falls below 10^{-8}.

Figure 3a plots gain in solid lines (left y-axis) and NF in dashed lines (right y-axis). The circle markers are for the case when NF is optimized, while the cross markers are for the solution that minimizes gain ripple. Clearly the best solution is that minimizing NF: the gain is higher and the noise figure lower. The NF in this case is fairly flat, ranging from 7-9 dB. When minimizing gain ripple, the gain is very flat with a ripple of only 0.017 dB around 20.4 dB. However, we see that the NF is high and has a large excursion, reaching 15 dB for the shortest wavelength. Since the minimal gains are very similar ($\sim$20 dB) for the two optimizations a final stage filter could flatten all gain to 19.7 dB and keep NF below 9 dB.

Figure 3b shows the mid-stage band pass filter found for each optimization. We see that reducing the ripple by applying greater losses in the middle stage comes at a substantial cost in NF. We returned to laboratory and implemented both optimized mid-stage filters with our waveshaper. The experimental validation is included in Fig. 3a as dotted lines, and matches well to our prediction. This simple example shows the potential of NN for amplifier design. More complex scenarios involving a larger set of varying parameters, such as pump powers and fiber length, as well as more complex objective functions could be supported by this fast computational model combined with efficient optimization techniques.

4. Conclusion

We trained a feed forward neural network (FFNN) with experimentally collected training set. We used an experimental test set at lengths not covered in training to assess the quality of NN predictions of gain and noise figure (NF) for the second stage of a two-stage amplifier. The predictions were accurate with mean error of 0.2 dB for gain and 0.09 dB for NF. We used our model for joint optimization of the gain and ripple (or noise figure) via choice of a mid-stage band pass filter using particle swarm optimization (PSO) optimization. The NN model has lower computational complexity than physic-based models, and high accuracy.

References

1. Hanlin Feng, Saber Jalilpiran, Frédéric Maes, Lixian Wang, and Sophie LaRochelle. Characterization of Giles parameters for extended L-band erbium-doped fibers. *JOSA B*, 39(7):1783–1791, 2022.
2. Hamed Rabbani, Saber Jalilpiran, Sophie LaRochelle, and Leslie A. Rusch. Neural networks for fiber amplifier design optimization using experimental training sets. In *CLEO*, pages SF3H–2, 2023.
3. Jiakai Yu, Shengxiang Zhu, Craig L. Gutterman, Gil Zussman, and Daniel C. Kilper. Machine-learning-based EDFA gain estimation. *Journal of Optical Communications and Networking*, 13(4):B83–B91, 2021.
4. L. Wang and et al. Low cost solution for super L-band fiber amplifier based on single-mode and multi-mode hybrid pumping scheme. In *2022 Optical Fiber Communications Conf. (OFC)*, pages 1–3, 2022.